A Proof of Threshold Saturation for Spatially-Coupled LDPC Codes on BMS Channels

Santhosh Kumar[†], Andrew J. Young[†], Nicolas Macris[‡], and Henry D. Pfister[†]

Department of Electrical and Computer Engineering, Texas A&M University[†]

School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne[‡]

Abstract—Low-density parity-check (LDPC) convolutional codes have been shown to exhibit excellent performance under low-complexity belief-propagation decoding [1], [2]. This phenomenon is now termed threshold saturation via spatial coupling. The underlying principle behind this appears to be very general and spatially-coupled (SC) codes have been successfully applied in numerous areas. Recently, SC regular LDPC codes have been proven to achieve capacity universally, over the class of binary memoryless symmetric (BMS) channels, under belief-propagation decoding [3], [4].

In [5], [6], potential functions are used to prove that the BP threshold of SC irregular LDPC ensembles saturates, for the binary erasure channel, to the conjectured MAP threshold (known as the Maxwell threshold) of the underlying irregular ensembles. In this paper, that proof technique is generalized to BMS channels, thereby extending some results of [4] to irregular LDPC ensembles. We also believe that this approach can be expanded to cover a wide class of graphical models whose message-passing rules are associated with a Bethe free energy.

Index Terms—convolutional LDPC codes, density evolution, entropy functional, potential functions, spatial coupling, threshold saturation.

I. INTRODUCTION

Low-density parity-check (LDPC) convolutional codes were introduced in [7] and shown to have outstanding performance under belief-propagation (BP) decoding in [1], [2], [8]. The fundamental principle behind this phenomenon is described by Kudekar, Richardson, and Urbanke in [3] and coined *threshold saturation* via spatial coupling. For the binary erasure channel (BEC), they prove that spatially coupling a collection of (d_v, d_c)-regular LDPC ensembles produces a new (nearly) (d_v, d_c)-regular ensemble whose BP threshold approaches the MAP threshold of the original ensemble. Recently, a proof of threshold saturation (to the "area threshold") has been given for (d_v, d_c)-regular LDPC ensembles on binary memoryless symmetric (BMS) channels when d_v/d_c is fixed and d_v, d_c are sufficiently large [4]. This result implies that SC-LDPC codes achieve capacity universally over the class of BMS channels because the "area threshold" of regular LDPC codes approaches the Shannon

limit uniformly over this class when d_v/d_c is fixed and d_v, d_c are increased.

The idea of threshold saturation via spatial coupling has recently started a small revolution in coding theory, and SC codes have now been observed to *universally* approach the capacity regions of many systems [8], [9], [10], [11], [12], [13], [14], [15]. For spatially-coupled systems with suboptimal component decoders, such as message-passing decoding of code-division multiple access (CDMA) [16], [17] or iterative hard-decision decoding of SC generalized LDPC codes [18], the threshold saturates instead to an intrinsic threshold defined by the suboptimal component decoders.

SC has also led to new results for K-SAT, graph coloring, and the Curie-Weiss model in statistical physics [19], [20], [21]. For compressive sensing, SC measurement matrices were introduced in [22], shown to give large improvements with Gaussian approximated BP reconstruction in [23], and finally proven to achieve the information-theoretic limit in [24]. Recent results based on spatial coupling are now too numerous to cite thoroughly.

A different proof technique, based on potential functions, is used in [5] to prove that the BP threshold of spatially-coupled (SC) irregular LDPC ensembles saturates to the conjectured MAP threshold (known as the Maxwell threshold) of the underlying irregular ensembles. This technique is closely related to the analysis in [19], [20] for the Curie-Weiss model, the heuristic approach in [25], and the continuum approach used to prove threshold saturation for compressed sensing in [24]. In this paper, the proof technique based on potential functions (in [5], [6]) is extended to BMS channels. This extends the results of [4] by proving threshold saturation to the Maxwell threshold for BMS channels and a wide class of SC irregular LDPC ensembles. The main result is summarized in the following theorem whose proof comprises the majority of this paper.

Theorem 1: Consider the SC LDPC(λ, ρ) ensemble defined in Section IV-A and a family of BMS channels c(h) that is ordered by degradation and parameterized by entropy, h. For any BMS channel c(h) with h less than the potential threshold h*, there exists a sufficiently large coupling parameter w_0 such that, for all $w > w_0$, the SC density evolution converges to the perfect decoding solution.

Many observations, formal proofs, and a large variety of applications systems bear evidence to the generality of

threshold saturation. In particular, the approach taken in this paper can be seen as analyzing the average Bethe free energy of the SC ensemble in the large system limit [26], [27]. Therefore, it is tempting to conjecture that this approach can be applied to more general graphical models by computing the average Bethe free energy of the corresponding SC system.

II. PRELIMINARIES

A. Measure Algebras

We call a Borel measure x symmetric if

$$\int_{-E} e^{-\alpha/2}\,\mathsf{x}(d\alpha) = \int_E e^{-\alpha/2}\,\mathsf{x}(d\alpha),$$

for all Borel sets $E \subset \overline{\mathbb{R}}$ where one of the integrals is finite. Let $\mathcal{M} = M(\overline{\mathbb{R}})$ be the space of finite signed symmetric Borel measures on the extended real numbers $\overline{\mathbb{R}}$. In this work, the primary interest is on convex combinations and differences of symmetric probability measures that inherit many of their properties from $\mathcal{M}$. Let $\mathcal{X} \subset \mathcal{M}$ be the convex subset of symmetric probability measures. Also, let $\mathcal{X}_\mathrm{d} \subset \mathcal{M}$ be the subset of differences of symmetric probability measures:

$$\mathcal{X}_\mathrm{d} \triangleq \{\mathsf{x}_1 - \mathsf{x}_2 \mid \mathsf{x}_1, \mathsf{x}_2 \in \mathcal{X}\}.$$

In the interest of notational consistency, x is reserved for both finite signed symmetric Borel measures and symmetric probability measures, and y, z denote differences of symmetric probability measures.

In this space, there are two important binary operators, $\circledast$ and $\boxplus$, that denote the variable-node and check-node density evolution operations for log-likelihood ratio (LLR) message distributions, respectively. The wildcard $*$ is used to represent either operator in statements that apply to both operations. For example, the shorthand x^{*n} is used to denote

$$\underbrace{\mathsf{x} * \cdots * \mathsf{x}}_{n}.$$

For a polynomial $\lambda(\alpha) = \sum_{n=0}^{\deg(\lambda)} \lambda_n \alpha^n$ with real coefficients, we define

$$\lambda^*(\mathsf{x}) \triangleq \sum_{n=0}^{\deg(\lambda)} \lambda_n \mathsf{x}^{*n}.$$

Now, we give an explicit integral characterization of the operators $\circledast$ and $\boxplus$. For $\mathsf{x}_1, \mathsf{x}_2 \in \mathcal{M}$, and any Borel set $E \subset \overline{\mathbb{R}}$, define

$$(\mathsf{x}_1 \circledast \mathsf{x}_2)(E) \triangleq \int \mathsf{x}_1(E - \alpha)\,\mathsf{x}_2(d\alpha),$$

$$(\mathsf{x}_1 \boxplus \mathsf{x}_2)(E) \triangleq \int \mathsf{x}_1\left(2\tanh^{-1}\left(\frac{\tanh(\frac{E}{2})}{\tanh(\frac{\alpha}{2})}\right)\right)\mathsf{x}_2(d\alpha).$$

Associativity, commutativity, and linearity follow. Therefore, for measures x_1, x_2, $\mathsf{x}_3 \in \mathcal{M}$ and scalars α_1, $\alpha_2 \in \mathbb{R}$,

$$\mathsf{x}_1 * (\mathsf{x}_2 * \mathsf{x}_3) = (\mathsf{x}_1 * \mathsf{x}_2) * \mathsf{x}_3,$$
$$\mathsf{x}_1 * \mathsf{x}_2 = \mathsf{x}_2 * \mathsf{x}_1,$$
$$(\alpha_1 \mathsf{x}_1 + \alpha_2 \mathsf{x}_2) * \mathsf{x}_3 = \alpha_1(\mathsf{x}_1 * \mathsf{x}_3) + \alpha_2(\mathsf{x}_2 * \mathsf{x}_3).$$

Moreover, the space of symmetric probability measures is closed under these binary operations [28]. In a more abstract sense, the measure space $\mathcal{M}$ along with the multiplication operator $*$ forms a commutative unital associative algebra and this algebraic structure is induced on the space of symmetric probability measures. There is also an intrinsic connection between the algebras defined by each operator and one consequence is the duality (or conservation) result in Proposition 4. For the multiplicative identities in these algebras, $\mathsf{e}_\circledast = \Delta_0$ and $\mathsf{e}_\boxplus = \Delta_\infty$, we define $\mathsf{x}^{*0} = \mathsf{e}_*$ and observe the following relationships under the dual operation:

$$\Delta_0 \boxplus \mathsf{x} = \Delta_0 \qquad\qquad \Delta_\infty \circledast \mathsf{x} = \Delta_\infty.$$

In general, however, these operators do not associate

$$\mathsf{x}_1 \circledast (\mathsf{x}_2 \boxplus \mathsf{x}_3) \neq (\mathsf{x}_1 \circledast \mathsf{x}_2) \boxplus \mathsf{x}_3$$
$$\mathsf{x}_1 \boxplus (\mathsf{x}_2 \circledast \mathsf{x}_3) \neq (\mathsf{x}_1 \boxplus \mathsf{x}_2) \circledast \mathsf{x}_3,$$

nor distribute

$$\mathsf{x}_1 \circledast (\mathsf{x}_2 \boxplus \mathsf{x}_3) \neq (\mathsf{x}_1 \circledast \mathsf{x}_2) \boxplus (\mathsf{x}_1 \circledast \mathsf{x}_3)$$
$$\mathsf{x}_1 \boxplus (\mathsf{x}_2 \circledast \mathsf{x}_3) \neq (\mathsf{x}_1 \boxplus \mathsf{x}_2) \circledast (\mathsf{x}_1 \boxplus \mathsf{x}_3).$$

B. Partial Ordering by Degradation

Degradation is an important concept that allows one to compare some LLR message distributions. The order imposed by degradation is indicative of relating probability measures through a communication channel [28, Definition 4.69].

Definition 2: For $\mathsf{x} \in \mathcal{X}$ and $f : [0,1] \to \mathbb{R}$, define

$$I_f(\mathsf{x}) \triangleq \int f\left(\left|\tanh\left(\tfrac{\alpha}{2}\right)\right|\right)\mathsf{x}(d\alpha).$$

For $\mathsf{x}_1, \mathsf{x}_2 \in \mathcal{X}$, x_1 is said to be *degraded* with respect to x_2 (denoted $\mathsf{x}_1 \succeq \mathsf{x}_2$), if $I_f(\mathsf{x}_1) \geq I_f(\mathsf{x}_2)$ for all concave non-increasing f. Furthermore, x_1 is said to be *strictly degraded* with respect to x_2 (denoted $\mathsf{x}_1 \succ \mathsf{x}_2$) if $\mathsf{x}_1 \succeq \mathsf{x}_2$ and there is a concave non-increasing f so that $I_f(\mathsf{x}_1) > I_f(\mathsf{x}_2)$.

Degradation defines a partial order, on the space of symmetric probability measures, with maximal element Δ_0 and minimal element Δ_∞. In this paper, the notation for real intervals is overloaded and, for example, a half-open interval of measures is denoted by

$$(\mathsf{x}_1, \mathsf{x}_2] \triangleq \{\mathsf{x}' \in \mathcal{X} \mid \mathsf{x}_1 \prec \mathsf{x}' \preceq \mathsf{x}_2\}.$$

This partial ordering is also preserved under the binary operations as follows.

Proposition 3 ([28, Lemma 4.80]): Let $\mathsf{x}_1, \mathsf{x}_2 \in \mathcal{X}$ be ordered by degradation $\mathsf{x}_1 \succeq \mathsf{x}_2$. For any $\mathsf{x}_3 \in \mathcal{X}$, one has

$$\mathsf{x}_1 * \mathsf{x}_3 \succeq \mathsf{x}_2 * \mathsf{x}_3.$$

This ordering is our primary tool in describing relative channel quality, and thresholds. For further information see [28, pp. 204-208].

C. Entropy Functional for Symmetric Measures

Entropy is a fundamental quantity in information theory and communication, and it is defined as the linear functional, $H : \mathcal{M} \to \mathbb{R}$, given by the integral

$$H(x) \triangleq \int \log_2 \left(1 + e^{-\alpha}\right) x(d\alpha).$$

The entropy functional is the primary functional used in our analysis. It exhibits an order under degradation and, for symmetric probability measures $x_1 \succ x_2$, one has

$$H(x_1) > H(x_2).$$

The restriction to symmetric probability measures ($x \in \mathcal{X}$) also implies the bound

$$0 \leq H(x) \leq 1.$$

The operators $\circledast$ and $\boxplus$ admit a number of relationships under the entropy functional. The following results will prove invaluable in the ensuing analysis. Proposition 4 provides an important conservation result (also known as the duality rule for entropy) and Proposition 5 extends this relation to encompass differences of symmetric probability measures.

Proposition 4 ([28, pp. 196]): For $x_1, x_2 \in \mathcal{X}$, one has

$$H(x_1 \circledast x_2) + H(x_1 \boxplus x_2) = H(x_1) + H(x_2). \tag{1}$$

Proposition 5: For $x_1, x_2, x_3, x_4 \in \mathcal{X}$, one finds that

$$H(x_1 \circledast (x_3 - x_4)) + H(x_1 \boxplus (x_3 - x_4)) = H(x_3 - x_4),$$

$$H((x_1 - x_2) \circledast (x_3 - x_4)) + H((x_1 - x_2) \boxplus (x_3 - x_4)) = 0.$$

Proof: Consider the LHS of the first equality,

$$H(x_1 \circledast (x_3 - x_4)) + H(x_1 \boxplus (x_3 - x_4))$$
$$= H(x_1 \circledast x_3) + H(x_1 \boxplus x_3) - H(x_1 \circledast x_4) - H(x_1 \boxplus x_4)$$
$$= H(x_1) + H(x_3) - H(x_1) - H(x_4) \quad \text{(Proposition 4)}$$
$$= H(x_3 - x_4).$$

The second equality follows by expanding the LHS and applying the first equality twice. ∎

Due to the symmetry of the measures the entropy functional has an equivalent series representation.

Proposition 6: If $x \in \mathcal{M}$, then

$$H(x) = x(\overline{\mathbb{R}}) - \sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} \int \left(\tanh \frac{\alpha}{2}\right)^{2k} x(d\alpha).$$

Proof: For a sketch of the proof, see [29, Lemma 3]. ∎
Define the linear functional

$$M_k(x) \triangleq \int \left(\tanh \frac{\alpha}{2}\right)^{2k} x(d\alpha).$$

Since $-x^{2k}$ is a concave decreasing function over the interval $[0, 1]$, for symmetric probability measures $x_1 \succeq x_2$, by Definition 2,

$$M_k(x_1) \leq M_k(x_2).$$

Moreover, for $x \in \mathcal{X}$,

$$0 \leq M_k(x) \leq 1.$$

It is also easy to see that the functional $M_k(\cdot)$ takes the following product form under the operator $\boxplus$,

$$M_k(x_1 \boxplus x_2) = M_k(x_1) M_k(x_2).$$

Also, if $y_1, y_2 \in \mathcal{X}_d$ are the differences of symmetric probability measures, then $y_1(\overline{\mathbb{R}}) = y_2(\overline{\mathbb{R}}) = (y_1 * y_2)(\overline{\mathbb{R}}) = 0$. This leads to the following result.

Proposition 7: For $y_1, y_2 \in \mathcal{X}_d$, the entropy functional reduces to

$$H(y_1) = -\sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} M_k(y_1),$$

$$H(y_1 \boxplus y_2) = -\sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} M_k(y_1) M_k(y_2).$$

In particular,

$$H(y_1 \boxplus y_1) = -\sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} M_k(y_1)^2 \leq 0.$$

with equality iff $y_1 = 0$ almost everywhere.
The above proposition implies an important upper bound on the entropy functional for differences of symmetric probability measures under *either* operator, $\circledast$ or $\boxplus$.

Proposition 8: If $x_1, x_1', x_2, x_3 \in \mathcal{X}$ with $x_1' \succeq x_1$, then

$$|H((x_1' - x_1) * (x_2 - x_3))| \leq H(x_1' - x_1).$$

Proof: We show the result for the operator $\boxplus$. The extension to $\circledast$ follows from the duality rule for entropy for differences of symmetric probability measures, Proposition 5. From Proposition 7

$$|H((x_1' - x_1) \boxplus (x_2 - x_3))|$$
$$\leq \sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} |M_k(x_1' - x_1)| \, |M_k(x_2 - x_3)|$$
$$\overset{(a)}{=} -\sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} M_k(x_1' - x_1) \, |M_k(x_2 - x_3)|,$$
$$\overset{(b)}{\leq} -\sum_{k=1}^{\infty} \frac{(\log 2)^{-1}}{2k(2k-1)} M_k(x_1' - x_1)$$
$$= H(x_1' - x_1),$$

where (a) follows from $M_k(x_1') \leq M_k(x_1)$ and (b) follows since $0 \leq M_k(x_2), M_k(x_3) \leq 1$. ∎

Corollary 9: For $x_1, x_1', x_2, x_3, x_4 \in \mathcal{X}$ with $x_1' \succeq x_1$, one has

$$|H((x_1' - x_1) * (x_2 - x_3) * x_4)| \leq H(x_1' - x_1).$$

Proof: This follows from Proposition 8 by replacing x_2, x_3 with $x_2 * x_4$ and $x_3 * x_4$, respectively. ∎

D. Directional Derivatives

The main result in this paper is derived using potential theory and differential relations. One can avoid the technical challenges of differentiation in the abstract space of measures by focusing on directional derivatives of functionals that map measures to real numbers.

Definition 10: Let $F : \mathcal{M} \to \mathbb{R}$ be a functional on $\mathcal{M}$. The *directional derivative* of F at x in the direction y is

$$\mathrm{d}_{\mathsf{x}} F(\mathsf{x})[\mathsf{y}] \triangleq \lim_{\delta \to 0} \frac{F(\mathsf{x} + \delta \mathsf{y}) - F(\mathsf{x})}{\delta},$$

whenever the limit exists.

This definition is naturally extended to higher-order directional derivatives using

$$\mathrm{d}_{\mathsf{x}}^n F(\mathsf{x})[\mathsf{y}_1, \ldots, \mathsf{y}_n] \triangleq \mathrm{d}_{\mathsf{x}} \left(\cdots \mathrm{d}_{\mathsf{x}} \left(\mathrm{d}_{\mathsf{x}} F(\mathsf{x})[\mathsf{y}_1] \right)[\mathsf{y}_2] \cdots \right)[\mathsf{y}_n],$$

and vectors of measures using, for $\underline{\mathsf{x}} = [\mathsf{x}_1, \cdots, \mathsf{x}_m]$,

$$\mathrm{d}_{\underline{\mathsf{x}}} F(\underline{\mathsf{x}})[\underline{\mathsf{y}}] \triangleq \lim_{\delta \to 0} \frac{F(\underline{\mathsf{x}} + \delta \underline{\mathsf{y}}) - F(\underline{\mathsf{x}})}{\delta},$$

whenever the limit exists. Similarly, we can define higher-order directional derivatives of functionals on vectors of measures.

The utility of directional derivatives for linear functionals is evident from the following Lemma.

Lemma 11: Let $F : \mathcal{M} \to \mathbb{R}$ be a linear functional and $*$ be an associative, commutative, and linear binary operator. Then, for $\mathsf{x}, \mathsf{y}, \mathsf{z} \in \mathcal{M}$, we have

$$\mathrm{d}_{\mathsf{x}} F(\mathsf{x}^{*n})[\mathsf{y}] = n F(\mathsf{x}^{*(n-1)} * \mathsf{y}),$$
$$\mathrm{d}_{\mathsf{x}}^2 F(\mathsf{x}^{*n})[\mathsf{y}, \mathsf{z}] = n(n-1) F\left(\mathsf{x}^{*(n-2)} * \mathsf{y} * \mathsf{z}\right).$$

Proof: A binary operation $*$ that is associative, commutative, and linear admits a binomial expansion of the form

$$(\mathsf{x} + \delta \mathsf{y})^{*n} = \sum_{i=0}^n \delta^i \binom{n}{i} \mathsf{x}^{*(n-i)} * \mathsf{y}^{*i}.$$

Then, the linearity of F implies that

$$F\left((\mathsf{x} + \delta \mathsf{y})^{*n}\right) - F\left(\mathsf{x}^{*n}\right)$$
$$= \delta n F(\mathsf{x}^{*(n-1)} * \mathsf{y}) + \sum_{i=2}^n \delta^i F(\mathsf{x}^{*(n-i)} * \mathsf{y}^{*i}).$$

Dividing by δ and taking the limit gives

$$\mathrm{d}_{\mathsf{x}} F(\mathsf{x}^{*n})[\mathsf{y}] = n F(\mathsf{x}^{*(n-1)} * \mathsf{y}).$$

An analogous argument shows the result for the second-order directional derivative. ∎

Remark 12: In general, applying Taylor's theorem to some mapping $T : \mathcal{X} \to \mathcal{X}$ requires advanced mathematical machinery. However, in our problem, the entropy functional and its interplay with the operators $\circledast$ and $\boxplus$ impose a polynomial structure on the functions of interest, obviating the need for Fréchet derivatives. Therefore, Taylor's theorem becomes quite simple for parameterized linear functionals $\phi : [0, 1] \to \mathbb{R}$ of the form

$$\phi(t) = F(\mathsf{x}_1 + t(\mathsf{x}_2 - \mathsf{x}_1)).$$

III. SINGLE SYSTEM

Let LDPC(λ, ρ) denote the LDPC ensemble with variable-node degree distribution λ and check-node degree distribution ρ. The edge-perspective degree distributions λ, ρ have an equivalent representation in terms of the node-perspective degree distributions L, R, namely,

$$\lambda(\alpha) = \frac{L'(\alpha)}{L'(1)}, \qquad \rho(\alpha) = \frac{R'(\alpha)}{R'(1)}.$$

This differential relationship is crucial in developing an appropriate potential functional for the LDPC(λ, ρ) ensemble.

Density evolution (DE) characterizes the asymptotic performance of the LDPC(λ, ρ) ensemble by describing the evolution of message distributions with iteration. For this ensemble, the DE update is compactly described by

$$\tilde{\mathsf{x}}^{(\ell+1)} = \mathsf{c} \circledast \lambda^{\circledast}(\rho^{\boxplus}(\tilde{\mathsf{x}}^{(\ell)})),$$

where $\tilde{\mathsf{x}}^{(\ell)}$ is the variable-node output distribution after ℓ iterations of message passing [30], [28].

We now develop the necessary definitions for the single-system potential framework. Included are the potential functional, fixed points, stationary points, the directional derivative of the potential functional, and thresholds.

Definition 13: Consider a family of BMS channels whose LLR distributions $\mathsf{c}(\mathsf{h}) : [0, 1] \to \mathcal{X}$ are ordered by degradation and parameterized by their entropy $\mathrm{H}(\mathsf{c}(\mathsf{h})) = \mathsf{h}$. The *BP threshold channel* of such a family is defined to be $\mathsf{c}^{\mathrm{BP}} \triangleq \mathsf{c}(\mathsf{h}^{\mathrm{BP}})$, where $\mathsf{h}^{\mathrm{BP}} \triangleq$

$$\sup \left\{ \mathsf{h} \in [0, 1] \mid \mathsf{x} \succ \mathsf{c}(\mathsf{h}) \circledast \lambda^{\circledast}(\rho^{\boxplus}(\mathsf{x})) \, \forall \mathsf{x} \in (\Delta_\infty, \Delta_0] \right\}.$$

Definition 14: For a family of BMS channels, $\mathsf{c}(\mathsf{h})$, the MAP threshold is given by $\mathsf{h}^{\mathrm{MAP}} \triangleq$

$$\inf \left\{ \mathsf{h} \in [0, 1] \mid \liminf_{n \to \infty} \tfrac{1}{n} \mathbb{E}\left[\mathrm{H}\left(X^n \mid Y^n(\mathsf{c}(\mathsf{h}))\right) \right] > 0 \right\}.$$

Definition 15: The *potential functional* (or the average Bethe free energy), $U : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, of the LDPC$(\lambda, \rho)$ ensemble is

$$U(\mathsf{x}; \mathsf{c}) \triangleq \frac{L'(1)}{R'(1)} \mathrm{H}\left(R^{\boxplus}(\mathsf{x})\right) + L'(1)\mathrm{H}\left(\rho^{\boxplus}(\mathsf{x})\right)$$
$$- L'(1)\mathrm{H}\left(\mathsf{x} \boxplus \rho^{\boxplus}(\mathsf{x})\right) - \mathrm{H}\left(\mathsf{c} \circledast L^{\circledast}\left(\rho^{\boxplus}(\mathsf{x})\right)\right).$$

Remark 16: This potential functional is essentially the negative of the replica-symmetric free energies calculated in [29], [26], [31]. When applied to the binary erasure channel, it is a constant multiple of the potential function defined in [5]. An example of $U(\mathsf{x}; \mathsf{c})$ is shown in Fig. 1.

Definition 17: The *fixed-point potential*, $P : \mathcal{X} \to \mathbb{R}$, of the LDPC$(\lambda, \rho)$ ensemble is

$$P(\mathsf{x}) \triangleq U(\mathsf{x}; \mathsf{f}(\mathsf{x})),$$

where $\mathsf{f} : \mathcal{X} \to \mathcal{X}$ satisfies $\mathsf{f}(\mathsf{x}) \circledast \lambda^{\circledast}\left(\rho^{\boxplus}(\mathsf{x})\right) = \mathsf{x}$.

Remark 18: For regular codes, one can use duality rule for entropy to show that the fixed-point potential defined above is exactly equal to the negative of the GEXIT integral functional denoted by A in [4, Lemma 26]. A similar relationship with EBP GEXIT integrals is believed to be true in general, but the authors are not aware of a proof for this.
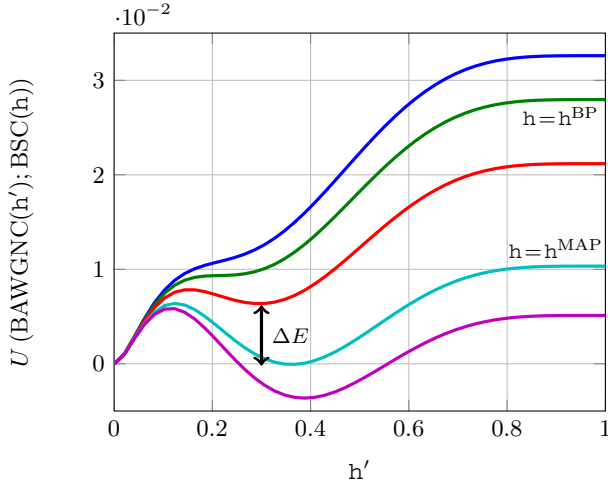
Fig. 1. Potential functional for $(\lambda, \rho) = (x^2, x^5)$ on the binary symmetric channel (BSC), with $h \in \{0.40, 0.416, 0.44, 0.469, 0.48\}$. The x-input is chosen to be the binary AWGN channel (BAWGNC) with entropy h'.

Definition 19: For $x \in \mathcal{X}$,

a) x is a *fixed point* of *density evolution* if

$$x = c \circledast \lambda^\circledast \left(\rho^\boxplus(x) \right),$$

b) x is a *stationary point* of the potential if, for all $y \in \mathcal{X}_d$,

$$d_x U(x; c)[y] = 0.$$

Lemma 20: For $x, c \in \mathcal{X}$ and $y \in \mathcal{X}_d$, the directional derivative of the potential functional with respect to x in the direction y is

$$d_x U(x; c)[y]$$
$$= L'(1) H \left(\left[x - c \circledast \lambda^\circledast \left(\rho^\boxplus(x) \right) \right] \circledast \left[\rho'^\boxplus(x) \boxplus y \right] \right).$$

Proof: The directional derivative for each of the four terms is calculated following the procedure outlined in the proof of Lemma 11. The first three terms are

$$d_x H \left(R^\boxplus(x) \right)[y] = R'(1) H \left(\rho^\boxplus(x) \boxplus y \right),$$
$$d_x H \left(\rho^\boxplus(x) \right)[y] = H \left(\rho'^\boxplus(x) \boxplus y \right),$$
$$d_x H \left(x \boxplus \rho^\boxplus(x) \right)[y] = H \left(\rho^\boxplus(x) \boxplus y \right) + H \left(x \boxplus \rho'^\boxplus(x) \boxplus y \right)$$
$$\overset{(a)}{=} H \left(\rho^\boxplus(x) \boxplus y \right) + H \left(\rho'^\boxplus(x) \boxplus y \right)$$
$$- H \left(x \circledast \left[\rho'^\boxplus(x) \boxplus y \right] \right),$$

where (a) follows from Proposition 5, $\rho'^\boxplus(x) \boxplus y$ is the difference of probability measures multiplied by the scalar $\rho'(1)$. Since the operators $\circledast$ and $\boxplus$ do not associate, one must exercise care in analyzing the last term,

$$d_x H \left(c \circledast L^\circledast \left(\rho^\boxplus(x) \right) \right)[y]$$
$$= L'(1) H \left(\left[c \circledast \lambda^\circledast(\rho^\boxplus(x)) \right] \circledast \left[\rho'^\boxplus(x) \boxplus y \right] \right).$$

Consolidating the four terms gives the desired result. ∎

Corollary 21: If $x \in \mathcal{X}$ is a fixed point of density evolution, then it is also a stationary point of the potential functional.

Definition 22: For a channel $c \in \mathcal{X}$, define the *basin of attraction* for Δ_∞ as

$$\mathcal{V}(c) \triangleq \left\{ x \in \mathcal{X} \mid x' \succ c \circledast \lambda^\circledast(\rho^\boxplus(x')) \; \forall x' \in (\Delta_\infty, x] \right\}.$$

Definition 23: For a channel $c \in \mathcal{X}$, the *energy gap* is defined as

$$\Delta E(c) \triangleq \inf_{x \in \mathcal{X} \backslash \mathcal{V}(c)} U(x; c).$$

Lemma 24: The following relations are important in characterizing the thresholds. Suppose $h_1 > h_2$. Then

a) $c(h_1) \succ c(h_2)$
b) $U(x; c(h_1)) < U(x; c(h_2))$, if $x \neq \Delta_\infty$.
c) $\mathcal{V}(c(h_1)) \subseteq \mathcal{V}(c(h_2)) \implies \mathcal{X} \backslash \mathcal{V}(c(h_1)) \supseteq \mathcal{X} \backslash \mathcal{V}(c(h_2))$
d) $\Delta E(c(h_1)) < \Delta E(c(h_2))$

Definition 25: The *potential threshold channel* of a family, $c(h)$, of BMS channels is given by $c^* \triangleq c(h^*)$, where

$$h^* \triangleq \sup\{h \in [h^{BP}, 1] \mid \Delta E(c(h)) > 0\}$$

Note that, if $h < h^*$, from Lemma 24d, $\Delta E(c(h)) > 0$.

Lemma 26: From [29], [26], [31], the following holds for degree distributions without odd-degree checks on any BMS channel and any degree distribution on the binary-input AWGN channel:

a) $\liminf_{n \to \infty} \frac{1}{n} \mathbb{E} \left[H \left(X^n | Y^n(c(h)) \right) \right] \geq - \inf_{x \in \mathcal{X}} U(x; c(h))$,
b) $h^{MAP} \leq h^*$.

Proof: Since the potential functional is the negative of the replica-symmetric free energies in [29], [26], [31], the main result of these papers translates directly into the first result. For the second result, consider any $h > h^*$. Below, we will arrive at a conclusion that $h \geq h^{MAP}$, which establishes the desired result. Note that from Lemma 24d, $\Delta E(c(h)) < 0$. From the first part of this lemma,

$$\liminf_{n \to \infty} \frac{1}{n} \mathbb{E} \left[H \left(X^n | Y^n(c(h)) \right) \right]$$
$$\geq - \inf_{x \in \mathcal{X}} U(x; c(h)) \geq - \inf_{x \in \mathcal{X} \backslash \mathcal{V}(c)} U(x; c(h))$$
$$= -\Delta(E(c(h))) > 0.$$

Hence, $h \geq h^{MAP}$. ∎

Remark 27: This implies that the MAP threshold is upper bounded by the potential threshold for the single system. It is conjectured that, in general, these two quantities are actually equal. Some progress has been made towards proving this for special cases [32].

IV. COUPLED SYSTEM

The potential theory from Section III is now extended to spatially-coupled systems. Vectors of measures are denoted by underlines (e.g., $\underline{x}$) with $[\underline{x}]_i = x_i$. Functionals operating on single densities are distinguished from those operating on vectors by their input (i.e., $U(x; c)$ vs. $U(\underline{x}; c)$). Also, for vectors $\underline{x}$ and $\underline{x}'$, we write $\underline{x} \succeq \underline{x}'$ if $x_i \succeq x_i'$ for all i, and $\underline{x} \succ \underline{x}'$ if $\underline{x} \succeq \underline{x}'$ and $x_i \succ x_i'$ for some i.

A. Spatial Coupling

The ideas underlying spatial coupling now appear to be quite general. The local coupling in the system allows the effect of the perfect information, provided at the boundary, to propagate throughout the system. In the large system limit, these coupled systems show a significant performance improvement.

The (λ, ρ, N, w) SC ensemble is defined as follows. A collection of $2N$ variable-node groups are placed at all positions in $\mathcal{N}_v = \{1, 2, \ldots, 2N\}$ and a collection of $2N + (w - 1)$ check-node groups are placed at all positions in $\mathcal{N}_c = \{1, 2, \ldots, 2N + (w - 1)\}$. For notational convenience, the rightmost check-node group index is denoted by $N_w \triangleq 2N + (w - 1)$. The integer M is chosen large enough so that i) ML_i, $ML'(1)R_j/R'(1)$ are natural numbers for $1 \le i \le \deg(L)$, $1 \le j \le \deg(R)$, and ii) $ML'(1)$ is divisible by w.

At each variable-node group, ML_i nodes of degree i are placed for $1 \le i \le \deg(L)$. Similarly, at each check-node group, $ML'(1)R_i/R'(1)$ nodes of degree i are placed for $1 \le i \le \deg(R)$. At each variable-node and check-node group the $ML'(1)$ sockets are partitioned into w equal-sized groups using a uniform random permutation. Denote these partitions, respectively, by $\mathcal{P}^v_{i_1,j}$ and $\mathcal{P}^c_{i_2,j}$ at variable-node and check-node groups, where $1 \le i_1 \le 2N$, $1 \le i_2 \le N_w$ and $1 \le j \le w$. The SC system is constructed by connecting the sockets from $\mathcal{P}^v_{i,j}$ to $\mathcal{P}^c_{i+j-1,j}$ using uniform random permutations. This construction leaves some sockets of the check-node groups at the boundaries unconnected and these sockets are assigned the binary value 0 (i.e., the socket and edge are removed). These 0 values form the seed that gets decoding started.

B. Spatially-Coupled Systems

Let $\tilde{\mathsf{x}}_i^{(\ell)}$ be the variable-node output distribution at node i after ℓ iterations of message passing. Then, the input distribution to the i-th check-node group is the normalized sum of averaged variable-node output distributions

$$\mathsf{x}_i^{(\ell)} = \frac{1}{w} \sum_{k=0}^{w-1} \tilde{\mathsf{x}}_{i-k}^{(\ell)}. \tag{2}$$

Backward averaging follows naturally from the setup and is essentially the transpose of the forward averaging for the check-node output distributions. This model uses uniform coupling over a fixed window, but in a more general setting window size and coefficient weights could vary from node to node. By virtue of the boundary conditions, $\tilde{\mathsf{x}}_i^{(\ell)} = \Delta_\infty$ for $i \notin \mathcal{N}_v$ and all ℓ, and from the relation in (2), this implies $\mathsf{x}_i^{(\ell)} = \Delta_\infty$ for $i \notin \mathcal{N}_c$ and all ℓ.

Generalizing [4, Eqn. 12] to irregular codes gives evolution of the variable-node output distributions,

$$\tilde{\mathsf{x}}_i^{(\ell+1)} = \mathsf{c} \circledast \lambda^{\circledast} \left(\frac{1}{w} \sum_{j=0}^{w-1} \rho^{\boxplus} \left(\frac{1}{w} \sum_{k=0}^{w-1} \tilde{\mathsf{x}}_{i+j-k}^{(\ell)} \right) \right). \tag{3}$$

Making a change of variables, the variable-node output distribution evolution in (3) can be rewritten in terms of check-node input distributions

$$\mathsf{x}_i^{(\ell+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \mathsf{c}_{i-k} \circledast \lambda^{\circledast} \left(\frac{1}{w} \sum_{j=0}^{w-1} \rho^{\boxplus} \left(\mathsf{x}_{i-k+j}^{(\ell)} \right) \right), \tag{4}$$

where $\mathsf{c}_i = \mathsf{c}$ when $i \in \mathcal{N}_v$ and $\mathsf{c}_i = \Delta_\infty$ otherwise. While (3) is a more natural representation for the underlying
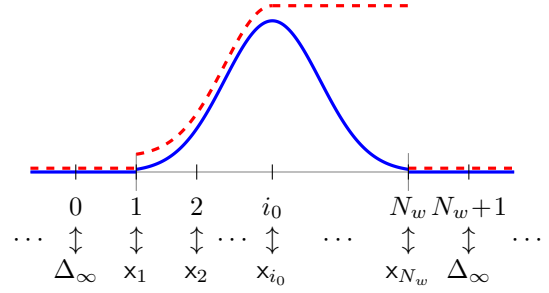

Fig. 2. This figure depicts the entropies of $\mathsf{x}_1, \cdots, \mathsf{x}_{N_w}$ in a typical iteration. The blue line (solid) corresponds to the two-sided system and the red line (dashed) to the one-sided system. The distributions of the one-sided system are always degraded with respect to the two-sided system, hence a higher entropy. The distributions outside the set $\{1, \cdots, N_w\}$ are fixed to Δ_∞ for both the systems.

system, (4) is more mathematically tractable and easily yields a coupled potential functional. As such, we adopt the system characterized by (4) and refer to it as the *two-sided spatially-coupled system*.

The two-sided spatially-coupled system is initialized with

$$\mathsf{x}_i^{(0)} = \Delta_0, \quad 1 \le i \le N_w.$$

This symmetric initialization, uniform coupling coefficients, and symmetric boundary conditions (i.e., seed information) induce symmetry on all the message distributions. In particular, the two-sided system is fully described by only half the distributions because

$$\mathsf{x}_i^{(\ell)} = \mathsf{x}_{2N+w-i}^{(\ell)},$$

for all ℓ. As density evolution progresses, the perfect boundary information from the left and right sides propagates inward. This propagation induces a nondecreasing degradation ordering on positions $1, \ldots, \lceil N_w/2 \rceil$ and a nonincreasing ordering on positions $\lceil (N_w + 1)/2 \rceil, \ldots, N_w$ [4, Def. 44].

The nondecreasing ordering by degradation introduces a degraded maximum at $i_0 \triangleq N + \lfloor \frac{w}{2} \rfloor$, and this maximum allows one to define a modified recursion that upper bounds the two-sided spatially-coupled system.

Definition 28 ([4]): The *one-sided spatially-coupled system* is a modification of (4) defined by fixing the values of positions outside $\mathcal{N}'_c \triangleq \{1, 2, \ldots, i_0\}$, where i_0 is defined as above. As before, the boundary is fixed to Δ_∞, that is $\mathsf{x}_i^{(\ell)} = \Delta_\infty$ for $i \notin \mathcal{N}_c$ and all ℓ. More importantly, it also fixes the values $\mathsf{x}_i^{(\ell)} = \mathsf{x}_{i_0}^{(\ell)}$ for $i_0 < i \le N_w$ and all ℓ.

The density evolution update for the one-sided system is identical to (4) for the first i_0 terms, $\{\mathsf{x}_1^{(\ell)}, \mathsf{x}_2^{(\ell)}, \ldots, \mathsf{x}_{i_0}^{(\ell)}\}$. But, for the remaining terms, it simply repeats the distribution $\mathsf{x}_{i_0}^{(\ell)}$ and this implies $\mathsf{x}_i^{(\ell)} = \mathsf{x}_{i_0}^{(\ell)}$ for $i_0 < i \le N_w$ at every step. The one-sided and two-sided systems are illustrated in Fig. 2.

Lemma 29: For the one-sided spatially-coupled system, the fixed point resulting from the density evolution satisfies

$$\mathsf{x}_i \succeq \mathsf{x}_{i-1}, \quad 1 \le i \le N_w$$

Proof: See [4, Section IV-D]. ∎

$$U(\underline{x}; c) \triangleq L'(1) \sum_{i=1}^{N_w} \left[\frac{1}{R'(1)} H\left(R^{\boxtimes}(x_i)\right) + H\left(\rho^{\boxtimes}(x_i)\right) - H\left(x_i \boxtimes \rho^{\boxtimes}(x_i)\right) \right] - \sum_{i=1}^{2N} H\left(c \circledast L^{\circledast}\left(\frac{1}{w} \sum_{j=0}^{w-1} \rho^{\boxtimes}(x_{i+j})\right) \right) \quad (5)$$

$$d_{\underline{x}} U(\underline{x}; c)[\underline{y}] = L'(1) \sum_{i=1}^{N_w} H\left(\left[x_i - \frac{1}{w} \sum_{k=0}^{w-1} c_{i-k} \circledast \lambda^{\circledast}\left(\frac{1}{w} \sum_{j=0}^{w-1} \rho^{\boxtimes}(x_{i-k+j}) \right) \right] \circledast \left[\rho'^{\boxtimes}(x_i) \boxtimes y_i \right] \right) \quad (6)$$

$$d_{\underline{x}}^2 U(\underline{x}; c)[\underline{y}, \underline{z}] =$$

$$L'(1)\rho''(1) \sum_{i=1}^{N_w} H\left(\left[\frac{1}{w} \sum_{k=0}^{w-1} c_{i-k} \circledast \lambda^{\circledast}\left(\frac{1}{w} \sum_{j=0}^{w-1} \rho^{\boxtimes}(x_{i-k+j}) \right) \circledast \frac{\rho''^{\boxtimes}(x_i)}{\rho''(1)} \right] \boxtimes y_i \boxtimes z_i \right)$$

$$- L'(1)\rho''(1) \sum_{i=1}^{N_w} H\left(\left[x_i \boxtimes \frac{\rho''^{\boxtimes}(x_i)}{\rho''(1)} \right] \boxtimes y_i \boxtimes z_i \right) - L'(1)\rho'(1) \sum_{i=1}^{N_w} H\left(\frac{\rho'^{\boxtimes}(x_i)}{\rho'(1)} \boxtimes y_i \boxtimes z_i \right) \quad (7)$$

$$- \frac{L'(1)\lambda'(1)\rho'(1)^2}{w} \sum_{i=1}^{N_w} \sum_{m=\max\{i-(w-1),1\}}^{\min\{i+(w-1),N_w\}} H\left(\frac{1}{w} \sum_{k=0}^{w-1} c_{i-k} \circledast \frac{\lambda'^{\circledast}\left(\frac{1}{w} \sum_{j=0}^{w-1} \rho^{\boxtimes}(x_{i-k+j})\right)}{\lambda'(1)} \circledast \left[\frac{\rho'^{\boxtimes}(x_m)}{\rho'(1)} \boxtimes z_m \right] \circledast \left[\frac{\rho'^{\boxtimes}(x_i)}{\rho'(1)} \boxtimes y_i \right] \right)$$

C. Spatially-Coupled Potential

Definition 30: The potential functional for a spatially-coupled system $U : \mathcal{X}^{N_w} \times \mathcal{X} \to \mathbb{R}$ is given by (5).

Lemma 31: For a spatially-coupled system, the directional derivative of the potential functional with respect to $\underline{x} \in \mathcal{X}^{N_w}$ evaluated in the direction $\underline{y} \in \mathcal{X}_d^{N_w}$ is given by (6).

  Proof: The proof is similar to the single system case and is omitted for brevity. ∎

Lemma 32: For a spatially-coupled system, the second-order directional derivative of the potential functional with respect to $\underline{x}$ evaluated in the direction $[\underline{y}, \underline{z}] \in \mathcal{X}_d^{N_w} \times \mathcal{X}_d^{N_w}$ is given by (7).

  Proof: We skip the proof for brevity. ∎

V. A Proof of Threshold Saturation

We now prove threshold saturation for the spatially-coupled LDPC(λ, ρ) ensemble. Consider a spatially-coupled system with potential functional $U : \mathcal{X}^{N_w} \times \mathcal{X} \to \mathbb{R}$ as in Definition 30, and a parameterization $\phi : [0, 1] \to \mathbb{R}$

$$\phi(t) = U(\underline{x} + t(\underline{x}' - \underline{x}); c),$$

where $\underline{x}$ is a non-trivial fixed point of the one-sided SC system resulting from DE. The path endpoint $\underline{x}'$ is chosen to impose a small perturbation on $\underline{x}$. For all channels better than the potential threshold, $c \prec c(h^*)$, and any non-trivial fixed point, it can be shown that the one-sided SC potential strictly decreases along this perturbation. However, since a fixed point is also a stationary point of the potential functional, all variations in the potential up to the second-order can be made arbitrarily small by choosing a large coupling parameter w. Thus, one obtains a contradiction to the existence of a non-trivial fixed point from the second-order Taylor expansion of $\phi(t)$.

These ideas are formalized below. A right shift is chosen for the perturbation. This shift operator is defined in Definition 33 and Lemmas 34, 35 discuss its effect on the

potential functional at this shift and the directional derivative along the shift direction $[\mathbf{S}(\underline{x}) - \underline{x}]$. Lemma 36 establishes an important bound on the second-order directional derivative of the potential. Theorem 37 proves threshold saturation.

Definition 33: The shift operator $\mathbf{S} : \mathcal{X}^{N_w} \to \mathcal{X}^{N_w}$ is defined pointwise by

$$[\mathbf{S}(\underline{x})]_1 \triangleq \Delta_\infty, \qquad [\mathbf{S}(\underline{x})]_i \triangleq x_{i-1}, \quad 2 \leq i \leq N_w.$$

Lemma 34 ([5, Lemma 4]): Let $\underline{x} \in \mathcal{X}^{N_w}$ be such that $x_i = x_{i_0}$, for $i_0 \leq i \leq N_w$. Then, the change in the potential functional for a spatially-coupled system associated with the shift operator is bounded by

$$U(\mathbf{S}(\underline{x}); c) - U(\underline{x}; c) \leq -U(x_{i_0}; c).$$

  Proof: Due to boundary conditions, $x_i = x_{i_0}$ for $i_0 \leq i \leq N_w$, the only terms that contribute to $U(\mathbf{S}(\underline{x}); c) - U(\underline{x}; c)$ are given in (8). As the last w values, that is x_i for $2N \leq i \leq N_w$, are equal to x_{i_0}, we have the desired result. ∎

Lemma 35: If $\underline{x} \succ \underline{\Delta_\infty} \triangleq (\Delta_\infty, \dots, \Delta_\infty)$ is a fixed point of the one-sided spatially-coupled system, then

$$d_{\underline{x}} U(\underline{x}; c)[\mathbf{S}(\underline{x}) - \underline{x}] = 0,$$

and $x_{i_0} \notin \mathcal{V}(c)$ (i.e., x_{i_0} not in basin of attraction of Δ_∞).

  Proof: See Appendix. ∎

Lemma 36: Let $\underline{x}_1 \in \mathcal{X}^{N_w}$ be a vector of symmetric probability measures, and let $\underline{x}_2 \in \mathcal{X}^{N_w}$ be a vector of coupled check-node inputs ordered by degradation, $[x_2]_i \succeq [x_2]_{i-1}$, generated by coupling a vector of variable-node outputs $\underline{\tilde{x}}_2 \in \mathcal{X}^{2N}$

$$[x_2]_i = \frac{1}{w} \sum_{k=0}^{w-1} [\tilde{x}_2]_{i-k}.$$

The second-order directional derivative of $U(\underline{x}_1; c)$ with respect to $\underline{x}_1$ evaluated along $[\mathbf{S}(\underline{x}_2) - \underline{x}_2, \mathbf{S}(\underline{x}_2) - \underline{x}_2]$ can be absolutely bounded with

$$\left| d_{\underline{x}_1}^2 U(\underline{x}_1; c)[\mathbf{S}(\underline{x}_2) - \underline{x}_2, \mathbf{S}(\underline{x}_2) - \underline{x}_2] \right| \leq \frac{K_{\lambda, \rho}}{w},$$

$$U(\mathbf{S}(\underline{x}); c) - U(\underline{x}; c) = -\frac{L'(1)}{R'(1)} \mathrm{H}\left(R^{\boxplus}(x_{N_w})\right) - L'(1)\mathrm{H}\left(\rho^{\boxplus}(x_{N_w})\right) + L'(1)\mathrm{H}\left(x_{N_w} \boxplus \rho'^{\boxplus}(x_{N_w})\right) \tag{8}$$

$$+ \mathrm{H}\left(c \circledast L^{\circledast}\Big(\frac{1}{w}\sum_{j=0}^{w-1}\rho^{\boxplus}(x_{2N+j})\Big)\right) - \mathrm{H}\left(c \circledast L^{\circledast}\Big(\frac{1}{w}\sum_{j=0}^{w-1}\rho^{\boxplus}(x_{1+j})\Big)\right)$$

where the constant

$$K_{\lambda,\rho} \triangleq \gamma L'(1)\left(2\rho''(1) + \rho'(1) + 2\lambda'(1)\rho'(1)^2\right)$$

is independent of N and w.

Proof: See Appendix. ∎

Theorem 37: Consider a family of BMS channels $c(h)$ that is ordered by degradation and parameterized by the entropy, h. For a spatially-coupled LDPC (λ, ρ) ensemble with a coupling window $w > K_{\lambda,\rho}/(2\Delta E(c(h)))$, and a channel $c(h)$ with $h < h^*$, the only fixed point of density evolution is $\underline{\Delta}_\infty$.

Proof: Consider a one-sided spatially-coupled system. Fix $w > K_{\lambda,\rho}/(2\Delta E(c(h)))$. Suppose $\underline{x} \succ \underline{\Delta}_\infty$ is a fixed point of density evolution. Let $\underline{y} = \mathbf{S}(\underline{x}) - \underline{x}$ and $\phi : [0,1] \to \mathbb{R}$ be defined by

$$\phi(t) = U(\underline{x} + t\underline{y}; c(h)),$$

it is important to note that, for all $t \in [0,1]$, $\underline{x} + t\underline{y} = (1-t)\underline{x} + t\mathbf{S}(\underline{x})$ is a vector of probability measures. By linearity of the entropy functional and binary operators $\circledast$ and $\boxplus$, ϕ is a polynomial in t, and thus infinitely differentiable over the entire unit interval. The second-order Taylor series expansion about $t = 0$ evaluated at $t = 1$ provides

$$\phi(1) = \phi(0) + \phi'(0)(1-0) + \tfrac{1}{2}\phi''(t_0)(1-0)^2, \tag{9}$$

for some $t_0 \in [0,1]$. The first and second derivatives of ϕ are characterized by the first- and second-order directional derivatives of U:

$$\phi'(t) = \lim_{\delta \to 0} \frac{U(\underline{x} + (t+\delta)\underline{y}; c(h)) - U(\underline{x} + t\underline{y}; c(h))}{\delta}$$
$$= \mathrm{d}_{\underline{x}+t\underline{y}} U(\underline{x} + t\underline{y}; c(h))[\underline{y}],$$

and similarly,

$$\phi''(t) = \mathrm{d}^2_{\underline{x}+t\underline{y}} U(\underline{x} + t\underline{y}; c(h))[\underline{y}, \underline{y}].$$

Substituting and rearranging terms in (9) provides

$$\tfrac{1}{2}\mathrm{d}^2_{\underline{x}+t_0\underline{y}} U(\underline{x} + t_0\underline{y}; c(h))[\underline{y}, \underline{y}]$$
$$= U(\mathbf{S}(\underline{x}); c(h)) - U(\underline{x}; c(h)) - \mathrm{d}_{\underline{x}} U(\underline{x}; c(h))[\mathbf{S}(\underline{x}) - \underline{x}]$$
$$= U(\mathbf{S}(\underline{x}); c(h)) - U(\underline{x}; c(h)) \quad \text{(Lemma 35)}$$
$$\leq -U(x_{i_0}; c(h)) \quad \text{(Lemma 34)}$$
$$\leq -\Delta E(c(h)). \quad \text{(Lemma 35 and Definition 22)}$$

Taking an absolute value and applying the second-order directional derivative bound from Lemma 36 gives

$$\Delta E(c(h)) \leq \frac{K_{\lambda,\rho}}{2w} \quad \implies \quad w \leq \frac{K_{\lambda,\rho}}{2\Delta E(c(h))},$$

a contradiction. Hence the only fixed point of the one-sided system is $\underline{\Delta}_\infty$. The densities of the one-sided system are

degraded with respect to the two-sided system, and therefore, the only fixed point of the two-sided system is also $\underline{\Delta}_\infty$. ∎

VI. Conclusions

In this paper, the proof technique based on potential functions (in [5], [6]) is extended to BMS channels. This extends the results of [4] by proving threshold saturation to the Maxwell threshold for BMS channels and the class of SC irregular LDPC ensembles. In particular, for any family of BMS channels $c(h)$ that is ordered by degradation and parameterized by entropy, h, a potential threshold h^* is defined that satisfies $h^{\mathrm{MAP}} \leq h^*$, where h^{MAP} defines the MAP threshold of the underlying LDPC(λ, ρ) ensemble.

The main result is that, for sufficiently large w, the SC DE equation converges to the perfect decoding fixed point for any BMS channel $c(h)$ with $h < h^*$. The approach taken in this paper can be seen as analyzing the average Bethe free energy of the SC ensemble in the large system limit [26], [27].

This result reiterates the generality of the threshold saturation phenomenon, which is now evident from the many observations and proofs that span a wide variety of systems. We also believe that this approach can be extended to more general graphical models by computing the average Bethe free energy of the corresponding SC system.

Appendix

In the appendix, for a vector of measures $\underline{x}_i$, $x_{i,j}$ is used to denote $[\underline{x}_i]_j$.

A. Proof of Lemma 35

Since $\underline{x}$ is a fixed point of the one-sided spatially coupled system,

$$x_i = \frac{1}{w}\sum_{k=0}^{w-1} c_{i-k} \circledast \lambda^{\circledast}\left(\frac{1}{w}\sum_{j=0}^{w-1}\rho^{\boxplus}(x_{i-k+j})\right)$$

for $1 \leq i \leq i_0$, and $[\mathbf{S}(\underline{x}) - \underline{x}]_i = 0$ for $i_0 < i \leq N_w$. The first result follows from applying these relations to the directional derivative, given in Lemma 31. Now, observe that

$$x_{i_0} = \frac{1}{w}\sum_{k=0}^{w-1} c_{i_0-k} \circledast \lambda^{\circledast}\left(\frac{1}{w}\sum_{j=0}^{w-1}\rho^{\boxplus}(x_{i_0-k+j})\right)$$
$$\preceq \frac{1}{w}\sum_{k=0}^{w-1} c_{i_0-k} \circledast \lambda^{\circledast}\left(\frac{1}{w}\sum_{j=0}^{w-1}\rho^{\boxplus}(x_{i_0})\right)$$
$$\preceq c \circledast \lambda^{\circledast}\left(\rho^{\boxplus}(x_{i_0})\right).$$

Hence $x_{i_0} \notin \mathcal{V}(c)$.

B. Proof of Lemma 36

Let $\underline{y} = \mathbf{S}(\underline{x}_2) - \underline{x}_2$, with componentwise decomposition

$$y_j = [\mathbf{S}(\underline{x}_2) - \underline{x}_2]_j = x_{2,j-1} - x_{2,j}$$
$$= \frac{1}{w}\sum_{k=0}^{w-1}\tilde{x}_{2,j-1-k} - \frac{1}{w}\sum_{k=0}^{w-1}\tilde{x}_{2,j-k} = \frac{1}{w}(\tilde{x}_{2,j-w} - \tilde{x}_{2,j}),$$

where $\tilde{x}_{2,j} = \Delta_\infty$ for $j < 1$. Referencing Lemma 32, the first three terms of the second-order directional derivative are of the form

$$\mathrm{H}\left(x_3 \boxtimes y_i \boxtimes y_i\right) = \frac{1}{w}\mathrm{H}\left(x_3 \boxtimes (\tilde{x}_{2,i} - \tilde{x}_{2,i-w}) \boxtimes (x_{2,i} - x_{2,i-1})\right),$$

by linearity. From Corollary 9, this term is absolutely bounded by

$$\left|\mathrm{H}\left(x_3 \boxtimes y_i \boxtimes y_i\right)\right| \leq \frac{1}{w}\mathrm{H}\left(x_{2,i} - x_{2,i-1}\right).$$

The final term is of the form

$$\left|\mathrm{H}\left(x_4 \circledast [x_5 \boxtimes y_m] \circledast [x_6 \boxtimes y_i]\right)\right|$$
$$= \left|\mathrm{H}\left([x_4 \circledast (x_5 \boxtimes y_m)] \boxtimes [x_6 \boxtimes y_i]\right)\right| \quad \text{(Proposition 5)}$$
$$= \left|\mathrm{H}\left(x_6 \boxtimes [x_4 \circledast (x_5 \boxtimes y_m)] \boxtimes y_i\right)\right|$$
$$= \frac{1}{w}\left|\mathrm{H}\left(x_6 \boxtimes [x_8 - x_7] \boxtimes [x_{2,i} - x_{2,i-1}]\right)\right|$$
$$\leq \frac{1}{w}\mathrm{H}\left(x_{2,i} - x_{2,i-1}\right). \quad \text{(Corollary 9)}$$

By telescoping, one observes

$$\sum_{i=1}^{N_w}\mathrm{H}\left(x_{2,i} - x_{2,i-1}\right) = \mathrm{H}\left(x_{2,N_w} - \Delta_\infty\right) \leq 1.$$

Consolidating the above observations provides the desired upper bound.

REFERENCES

[1] A. Sridharan, M. Lentmaier, D. J. Costello, and K. S. Zigangirov, "Convergence analysis of a class of LDPC convolutional codes for the erasure channel," in *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, (Monticello, IL), pp. 953–962, 2004.

[2] M. Lentmaier, A. Sridharan, K. S. Zigangirov, and D. J. Costello, "Terminated LDPC convolutional codes with thresholds close to capacity," in *Proc. IEEE Int. Symp. Inform. Theory*, (Adelaide, Australia), pp. 1372–1376, 2005.

[3] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 803–834, 2011.

[4] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation." Arxiv preprint arXiv:1201.2999, 2012.

[5] A. Yedla, Y.-Y. Jian, P. S. Nguyen, and H. D. Pfister, "A simple proof of threshold saturation for coupled scalar recursions," in *Proc. Int. Symp. on Turbo Codes & Iterative Inform. Proc.*, pp. 51–55, 2012. Arxiv preprint arXiv:1204.5703, 2012.

[6] A. Yedla, Y.-Y. Jian, P. S. Nguyen, and H. D. Pfister, "A simple proof of threshold saturation for coupled vector recursions," in *Proc. IEEE Inform. Theory Workshop*, pp. 25–29, 2012. Arxiv preprint arXiv:1208.4080, 2012.

[7] J. Felstrom and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2181–2191, 1999.

[8] M. Lentmaier, A. Sridharan, D. J. Costello, and K. S. Zigangirov, "Iterative decoding threshold analysis for LDPC convolutional codes," *IEEE Trans. Inform. Theory*, vol. 56, pp. 5274–5289, Oct. 2010.

[9] S. Kudekar, C. Méasson, T. Richardson, and R. Urbanke, "Threshold saturation on BMS channels via spatial coupling," in *Proc. Int. Symp. on Turbo Codes & Iterative Inform. Proc.*, pp. 309–313, Sept. 2010.

[10] V. Rathi, R. Urbanke, M. Andersson, and M. Skoglund, "Rate-equivocation optimally spatially coupled LDPC codes for the BEC wiretap channel," in *Proc. IEEE Int. Symp. Inform. Theory*, (St. Petersburg, Russia), pp. 2393–2397, July 2011.

[11] A. Yedla, H. D. Pfister, and K. R. Narayanan, "Universality for the noisy Slepian-Wolf problem via spatial coupling," in *Proc. IEEE Int. Symp. Inform. Theory*, (St. Petersburg, Russia), pp. 2567–2571, July 2011.

[12] S. Kudekar and K. Kasai, "Threshold saturation on channels with memory via spatial coupling," in *Proc. IEEE Int. Symp. Inform. Theory*, (St. Petersburg, Russia), pp. 2562–2566, July 2011.

[13] S. Kudekar and K. Kasai, "Spatially coupled codes over the multiple access channel," in *Proc. IEEE Int. Symp. Inform. Theory*, (St. Petersburg, Russia), pp. 2816–2820, July 2011.

[14] P. S. Nguyen, A. Yedla, H. D. Pfister, and K. R. Narayanan, "Spatially-coupled codes and threshold saturation on intersymbol-interference channels." to be submitted to *IEEE Trans. on Inform. Theory*, [Online]. Available: http://arxiv.org/abs/1107.3253, 2012.

[15] P. S. Nguyen, A. Yedla, H. D. Pfister, and K. R. Narayanan, "Threshold saturation of spatially-coupled codes on intersymbol-interference channels," in *Proc. IEEE Int. Conf. Commun.*, (Ottawa, Canada), pp. 2209–2214, June 2012.

[16] K. Takeuchi, T. Tanaka, and T. Kawabata, "Improvement of BP-based CDMA multiuser detection by spatial coupling," in *Proc. IEEE Int. Symp. Inform. Theory*, (St. Petersburg, Russia), pp. 1489–1493, July 2011.

[17] C. Schlegel and D. Truhachev, "Multiple access demodulation in the lifted signal graph with spatial coupling," in *Proc. IEEE Int. Symp. Inform. Theory*, (St. Petersburg, Russia), pp. 2989–2993, July 2011.

[18] Y.-Y. Jian, H. D. Pfister, and K. R. Narayanan, "Approaching capacity at high rates with iterative hard-decision decoding," in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 2696–2700, 2012.

[19] S. H. Hassani, N. Macris, and R. Urbanke, "Coupled graphical models and their thresholds," in *Proc. IEEE Inform. Theory Workshop*, (Dublin, Ireland), pp. 1–5, 2010.

[20] S. H. Hassani, N. Macris, and R. Urbanke, "Chains of mean-field models," *J. Stat. Mech.*, p. P02011, 2012.

[21] S. H. Hassani, N. Macris, and R. Urbanke, "Threshold saturation in spatially coupled constraint satisfaction problems." Arxiv preprint arXiv:1112.6320, 2011.

[22] S. Kudekar and H. D. Pfister, "The effect of spatial coupling on compressive sensing," in *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, (Monticello, IL), pp. 347–353, Oct. 2010.

[23] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová, "Statistical physics-based reconstruction in compressed sensing." Arxiv preprint arXiv:1109.4424, Sept. 2011.

[24] D. L. Donoho, A. Javanmard, and A. Montanari, "Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing." Arxiv preprint arXiv:1112.0708, Dec. 2011.

[25] K. Takeuchi, T. Tanaka, and T. Kawabata, "A phenomenological study on threshold improvement via spatial coupling," *IEICE Transactions*, vol. 95-A, no. 5, pp. 974–977, 2012.

[26] N. Macris, "Griffith–Kelly–Sherman correlation inequalities: A useful tool in the theory of error correcting codes," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 664–683, 2007.

[27] R. Mori, "Connection between annealed free energy and belief propagation on random factor graph ensembles," in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 2010–2014, 2011.

[28] T. J. Richardson and R. L. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[29] A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3221–3246, 2005.

[30] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599–618, Feb. 2001.

[31] S. Kudekar and N. Macris, "Sharp bounds for optimal decoding of low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 55, no. 10, pp. 4635–4650, 2009.

[32] A. Giurgiu, N. Macris, and R. Urbanke, "How to prove the Maxwell conjecture via spatial coupling - A proof of concept," in *Proc. IEEE Int. Symp. Inform. Theory*, 2012.